

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number  
WO 01/79468 A2

(51) International Patent Classification<sup>7</sup>: C12N 9/00

(21) International Application Number: PCT/US01/11869

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/197,590	13 April 2000 (13.04.2000)	US
60/198,403	19 April 2000 (19.04.2000)	US
60/200,185	28 April 2000 (28.04.2000)	US
60/202,234	5 May 2000 (05.05.2000)	US
60/203,509	11 May 2000 (11.05.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/198,403 (CIP)
Filed on	19 April 2000 (19.04.2000)
US	60/197,590 (CIP)
Filed on	13 April 2000 (13.04.2000)
US	60/200,185 (CIP)
Filed on	28 April 2000 (28.04.2000)
US	60/202,234 (CIP)
Filed on	5 May 2000 (05.05.2000)
US	60/203,509 (CIP)
Filed on	11 May 2000 (11.05.2000)

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). HAFALIA, April [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). BURFORD, Nell [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). RING, Huijun, Z. [US/US]; 625 Orange Avenue, Los Altos, CA 94022 (US). LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US). YAO, Monique, G. [US/US];

111 Frederick Court, Mountain View, CA 94043 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). DAS, Debopriya [IN/US]; Apartment 3, 1179 Bonita Avenue, Mountain View, CA 94040 (US). SANJANWALA, Madhu, S. [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). GANDHI, Ameena, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). REDDY, Roopa [US/US]; 1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086 (US). KHAN, Farrah, A. [IN/US]; 333 Escuela Avenue, #221, Mountain View, CA 94040 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). RAMKUMAR, Jaya [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). GRIFFIN, Jennifer, A. [US/US]; 181 Irene Court, #11, Belmont, CA 94002 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DRUG METABOLIZING ENZYMES

(57) Abstract: The invention provides human drug metabolizing enzymes (DME) and polynucleotides which identify and encode DME. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of DME.

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	2434655CD1	527	S3 T120 T169 T200 S435 T520 S70 S343 S62 T141 T204 T243 T258 S296 S297 T419 Y130 Y143	N131 N313 N518	Signal peptide: M1-C23 Signal cleavage: M1-C23 Transmembrane domains: V91-N111, I492-F512 UDP-glucoronosyltransferase: G24-K525 UDP-glycosyltransferases BL00375: A: S34-L56; B: C125-P165; C: P188-N211; D: I253-C280; E: F293-P342; F: N348-P392; G: H447-H486 UDP-glucoronosyl and UDP-glucosyl transferases signature: N376-T417 UDP-glucoronosyltransferase (PD000190): G24-K338, V386-E527 UDP-glucoronosyl and glucosyl transferases: DM00367 P36510 176-459: M178-F460 UDP-glucoronosyltransferase signature: W354-Q397	HMMER SPScan HMMER HMMER-PFAM BLIMPS-BLOCKS
2	2516747CD1	523	S107 S81 S95 T99 S126 S213 S219 T224 S374 S59 S133 S362 S409 T423 Y74 Y433	N52	Signal peptide: M1-A22 Signal cleavage: M1-L19 transmembrane domain: Y484-C503 UDP-glucoronosyl and UDP-glucosyl transferase: A23-K521 UDP-glucoronosyl and UDP-glucosyl transferases signature: N373-T416 UDP-glycosyltransferases BL00375: A: S34-L56; B: C121-P161; C: P180-N203; D: L246-E273; F: N345-P389; G: L444-Q483 UDP-glucoronosyltransferase: PD000190: K24-G291, L246-D430, L422-K522 UDP-glucoronosyl and UDP-glucosyl transferases: DM00367 P36513 188-462: L177-V456	HMMER SPScan HMMER HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLAST-PRODOR MOTIFS HMMER SPScan HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLAST-PRODOR BLAST-DOMO

using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser. 7*:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser. 7*:225-232.) Alternatively, DME itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques.

- 5 (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of DME, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide
- 10 or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

- 15 In order to express a biologically active DME, the nucleotide sequences encoding DME or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in
- 20 polynucleotide sequences encoding DME. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding DME. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding DME and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional
- 25 or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See,
- 30 e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

- Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding DME and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989)
- 35 Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and

16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding DME. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding DME. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding DME can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding DME into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of DME are needed, e.g. for the production of antibodies, vectors which direct high level expression of DME may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of DME. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such

vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

5 Plant systems may also be used for expression of DME. Transcription of sequences encoding DME may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. 10 (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases 15 where an adenovirus is used as an expression vector, sequences encoding DME may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses DME in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma 20 virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino 25 polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of DME in cell lines is preferred. For example, sequences encoding DME can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous 30 expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be 35 propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding DME is inserted within a marker gene sequence, transformed cells containing sequences encoding DME can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding DME under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding DME and that express DME may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of DME using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on DME is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul

<400>	11						
gatcagtgtg	tgagggaaact	gccatcatga	ggtctgacaa	gtcagctttg	gtattttctgc	60	
tcctgcagct	cttctgtgtt	ggctgtggat	tctgtgggaa	agtcctgggt	tggccctgtg	120	
acatgagcca	ttggcttaat	gtcaaggctc	ttctagaaga	gctcatagtg	agaggccatg	180	
aggttaacagt	attgactcac	tcaaagcctt	cgtaaatga	ctacaggaag	ccttctgcac	240	
tgaattttga	ggtggtccat	atgccacagg	acagaacaga	agaaaatgaa	atatttgttg	300	
acctagctct	gaatgtcttg	ccaggcttat	caacctggca	atcagttata	aaattaaatg	360	
atTTTTTTgt	tgaataaaga	ggaactttaa	aatgatgtg	tgagagcttt	atctacaatc	420	
agagccttat	gaagaagcta	caggaaaacca	actacgatg	aatgcttata	gaccttgta	480	
ttccctgtgg	agacctgatg	gctgagttgc	ttgcagctcc	ttttgtgctc	acacttagaa	540	
tttctgtatg	aggcaatatg	gagcgaagct	gtgggaaact	tccagctcca	ctttcctatg	600	
tacctgtgcc	tatgacagga	ctaacagaca	gaatgacctt	tctggaaaga	gtaaaaaatt	660	
caatgctttc	agttttgttc	cacttctgga	ttcaggatta	cgactatcat	ttttgggaag	720	
agttttatag	taaggcatta	ggaaggccca	ctacattatg	tgagactatg	ggaaaagctg	780	
agatatggct	aatacgaaca	tattgggatt	ttgaatttcc	tcaacctata	caacctgaact	840	
ttgagtttgt	tggaggattg	cactgtaaac	ctgccaaagc	tttgccctaag	gaaatggaaa	900	
atTTTtTcca	gagttcaggg	gaagatggta	ttgtgggtgt	ttctctgggg	tactgttttc	960	
aaaatgttac	agaagaaaag	gctaatatca	ttgcttcagc	ccttgcccag	atcccacaga	1020	
agggtgttat	gaggtacaaa	ggaaaaaaac	catccacatt	aggagccaat	actcggctgt	1080	
atgatttgat	acccagaaat	gatcttcttg	gtcatcccaa	aaccaaaagt	tttatcactc	1140	
atggttggaat	gaatgggatc	tatgaagcta	tttaccatgg	ggtccttatg	gtgggaattc	1200	

```

ccatatttgg tgatcagctt gataacatag ctcacatgaa ggccaaagga gcagctgtag 1260
aaataaaactt caaaactatg acaagcgaag atttactgag ggcttttgaga acagtcatta 1320
ccgattcctc ttataaagag aatgctatga gattatcaag aattcaccat gatcaacctg 1380
taaagccccct agatcgagca gtcttctgga tcgagtttgt catgcgccac aaaggagcca 1440
agcacctgcg atcagctgcc catgacctca cctgggttcca gcactactct atagatgtga 1500
ttgggttccct gctgacctgt gtggcaactg ctatattctt gttcacaaaa tgttttttat 1560
tttctgtca aaaatttaat aaaactagaa agatagaaaa gagggaatag atctttccaa 1620
attcaagaaa gacctg                                     1636

```

&lt;210&gt; 12

&lt;211&gt; 2086

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2516747CB1

&lt;400&gt; 12

```

cctacctctt cctaggccca cagccagtgc ctttggagta ctgaggcgcg cacagagtcc 60
ttagcccgcc gcagggcgcg cagcccaggc tgagatccgc tgcttctgtg gaagtggagca 120
tgggtgggca gcgggtgctg cttctagtgg ccttccttct ttctggggtc ctgctctcag 180
aggctgccaa aatcctgaca atatctacac tgggtggaag ccattaccta ctgttggacc 240
gggtgtctca gattcttcaa gagcatggtc ataatgtgac tatgcttcat cagagtggaa 300
agtttttgat ccagatatt aaagaggagg aaaaatcata ccaagttatc agtggtttt 360
cacctgaaga tcatcaaaaa agaattaaga agcattttga tagctacata gaaacagcat 420
tggatggcag aaaagaatct gaagcccttg taaagctaag ggaaatattt gggactcaat 480
ttagttattt gctaagcaga aaggatataa tggattcctt aaagaatgag aactatgatc 540
tggattttgt tgaagcattt gatttctgtt ctttcctgat tgctgagaag cttgtgaaac 600
catttgtggc cattcttccc accacattcg gctctttgga ttttgggcta ccaagcccct 660
tgtcttatgt tccagtattc ccttccttgc tgactgatca catggacttc tggggccgag 720
tgaagaattt tctgatgttc tttagtttct ccaggagcca atgggacatg cagtctacat 780
ttgacaacac catcaaggag catttcccag aaggctctag gccagttttg tctcatcttc 840
tactgaaagc agagtgtgtg tttgttaact ctgattttgc ctttgatttt gcccgcccc 900
tgcttcccaa cactgtttat attggaggct tgatggaaaa acctattaaa ccagtaccac 960
aagacttggg caacttcatt gccactttg gggatgcagg gtttgtcctt gtggcctttg 1020
gtcccatgtt gaacacccat cagtcccagg aagtcctcaa gaagatgcac aatgcctttg 1080
cccacctccc tcaaggagtg atatggacat gtcagagttc tcattggccc agagatgttc 1140
atttggccac aaatgtgaaa attgtggact ggcttcctcg gagtgcctc ctggctcacc 1200
ccagcatccg tctttttgtc actcatggtg ggcagaacag cgtaatggag gccatccgtc 1260
atgggtgtgc catgggtgga ttaccagtca atggagacca gcatggaaac atggtcggag 1320
tagtagccaa aaattatggg gtctctatcc ggttgaatca ggtcacagcc gacacactga 1380
cacttacaat gaaacaagtc atagaagaca agaggtacaa gtcggcagtg gtggcagcca 1440
gtgtcatcct gcactctcag cccctgagcc ccgcacagcg gctgggtggg tggatcgacc 1500
acatcctcca gactggggga gcgacgcacc tcaagcccta tgccttccag cagccttggc 1560
atgagcagta cctcattgat gtcttttgtt ttctgctggg gctcactctg ggcactatgt 1620
ggctttgttg gaagctgctg ggtgtgtgtg ccagggtggc gcgtggggcc aggaaggtga 1680
agaagacatg aggctaggtg tagccttggg tgaggggagg gcatccctgg tcctttgaag 1740
gttctcccca cccagcaca cgccaccctt ctgttctctc ttcagctcca cctgccactg 1800
atcctgcaac ttgtctctt ctattctctg cctctgttta gaaatcttca cacaccactg 1860
aggcttcttg acttgcccct tgtgacttga attcccagct cagatacaaa ttttcacctg 1920
ccagccctgc ctctctctt ctcccttttc ctagacacag gactctgaca acttcatcct 1980
ccttgtttag atgacttccc agtttccagt cccatttctt ccttctatca cttttcataa 2040
aaaaactcag gaaatatttg acatatcttc catttcaaat tcttcc                                     2086

```

&lt;210&gt; 13

&lt;211&gt; 1814

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7472775CB1



Lys	Lys	Leu	Gln	Glu	Thr	Asn	Tyr	Asp	Val	Met	Leu	Ile	Asp	Pro
				140					145					150
Val	Ile	Pro	Cys	Gly	Asp	Leu	Met	Ala	Glu	Leu	Leu	Ala	Val	Pro
				155					160					165
Phe	Val	Leu	Thr	Leu	Arg	Ile	Ser	Val	Gly	Gly	Asn	Met	Glu	Arg
				170					175					180
Ser	Cys	Gly	Lys	Leu	Pro	Ala	Pro	Leu	Ser	Tyr	Val	Pro	Val	Pro
				185					190					195
Met	Thr	Gly	Leu	Thr	Asp	Arg	Met	Thr	Phe	Leu	Glu	Arg	Val	Lys
				200					205					210
Asn	Ser	Met	Leu	Ser	Val	Leu	Phe	His	Phe	Trp	Ile	Gln	Asp	Tyr
				215					220					225
Asp	Tyr	His	Phe	Trp	Glu	Glu	Phe	Tyr	Ser	Lys	Ala	Leu	Gly	Arg
				230					235					240
Pro	Thr	Thr	Leu	Cys	Glu	Thr	Val	Gly	Lys	Ala	Glu	Ile	Trp	Leu
				245					250					255
Ile	Arg	Thr	Tyr	Trp	Asp	Phe	Glu	Phe	Pro	Gln	Pro	Tyr	Gln	Pro
				260					265					270
Asn	Phe	Glu	Phe	Val	Gly	Gly	Leu	His	Cys	Lys	Pro	Ala	Lys	Ala
				275					280					285
Leu	Pro	Lys	Glu	Met	Glu	Asn	Phe	Val	Gln	Ser	Ser	Gly	Glu	Asp
				290					295					300
Gly	Ile	Val	Val	Phe	Ser	Leu	Gly	Ser	Leu	Phe	Gln	Asn	Val	Thr
				305					310					315
Glu	Glu	Lys	Ala	Asn	Ile	Ile	Ala	Ser	Ala	Leu	Ala	Gln	Ile	Pro
				320					325					330
Gln	Lys	Val	Leu	Trp	Arg	Tyr	Lys	Gly	Lys	Lys	Pro	Ser	Thr	Leu
				335					340					345
Gly	Ala	Asn	Thr	Arg	Leu	Tyr	Asp	Trp	Ile	Pro	Gln	Asn	Asp	Leu
				350					355					360
Leu	Gly	His	Pro	Lys	Thr	Lys	Ala	Phe	Ile	Thr	His	Gly	Gly	Met
				365					370					375
Asn	Gly	Ile	Tyr	Glu	Ala	Ile	Tyr	His	Gly	Val	Pro	Met	Val	Gly
				380					385					390
Val	Pro	Ile	Phe	Gly	Asp	Gln	Leu	Asp	Asn	Ile	Ala	His	Met	Lys
				395					400					405
Ala	Lys	Gly	Ala	Ala	Val	Glu	Ile	Asn	Phe	Lys	Thr	Met	Thr	Ser
				410					415					420
Glu	Asp	Leu	Leu	Arg	Ala	Leu	Arg	Thr	Val	Ile	Thr	Asp	Ser	Ser
				425					430					435
Tyr	Lys	Glu	Asn	Ala	Met	Arg	Leu	Ser	Arg	Ile	His	His	Asp	Gln
				440					445					450
Pro	Val	Lys	Pro	Leu	Asp	Arg	Ala	Val	Phe	Trp	Ile	Glu	Phe	Val
				455					460					465
Met	Arg	His	Lys	Gly	Ala	Lys	His	Leu	Arg	Ser	Ala	Ala	His	Asp
				470					475					480
Leu	Thr	Trp	Phe	Gln	His	Tyr	Ser	Ile	Asp	Val	Ile	Gly	Phe	Leu
				485					490					495
Leu	Thr	Cys	Val	Ala	Thr	Ala	Ile	Phe	Leu	Phe	Thr	Lys	Cys	Phe
				500					505					510
Leu	Phe	Ser	Cys	Gln	Lys	Phe	Asn	Lys	Thr	Arg	Lys	Ile	Glu	Lys
				515					520					525
Arg	Glu													

&lt;210&gt; 2

&lt;211&gt; 523

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2516747CD1

&lt;400&gt; 2

Met	Val	Gly	Gln	Arg	Val	Leu	Leu	Leu	Val	Ala	Phe	Leu	Leu	Ser	1	5	10	15
Gly	Val	Leu	Leu	Ser	Glu	Ala	Ala	Lys	Ile	Leu	Thr	Ile	Ser	Thr	20	25	30	35
Leu	Gly	Gly	Ser	His	Tyr	Leu	Leu	Leu	Asp	Arg	Val	Ser	Gln	Ile	40	45	50	55
Leu	Gln	Glu	His	Gly	His	Asn	Val	Thr	Met	Leu	His	Gln	Ser	Gly	60	65	70	75
Lys	Phe	Leu	Ile	Pro	Asp	Ile	Lys	Glu	Glu	Glu	Lys	Ser	Tyr	Gln	80	85	90	95
Val	Ile	Arg	Trp	Phe	Ser	Pro	Glu	Asp	His	Gln	Lys	Arg	Ile	Lys	100	105	110	115
Lys	His	Phe	Asp	Ser	Tyr	Ile	Glu	Thr	Ala	Leu	Asp	Gly	Arg	Lys	120	125	130	135
Glu	Ser	Glu	Ala	Leu	Val	Lys	Leu	Met	Glu	Ile	Phe	Gly	Thr	Gln	140	145	150	155
Cys	Ser	Tyr	Leu	Leu	Ser	Arg	Lys	Asp	Ile	Met	Asp	Ser	Leu	Lys	160	165	170	175
Asn	Glu	Asn	Tyr	Asp	Leu	Val	Phe	Val	Glu	Ala	Phe	Asp	Phe	Cys	180	185	190	195
Ser	Phe	Leu	Ile	Ala	Glu	Lys	Leu	Val	Lys	Pro	Phe	Val	Ala	Ile	200	205	210	215
Leu	Pro	Thr	Thr	Phe	Gly	Ser	Leu	Asp	Phe	Gly	Leu	Pro	Ser	Pro	220	225	230	235
Leu	Ser	Tyr	Val	Pro	Val	Phe	Pro	Ser	Leu	Leu	Thr	Asp	His	Met	240	245	250	255
Asp	Phe	Trp	Gly	Arg	Val	Lys	Asn	Phe	Leu	Met	Phe	Phe	Ser	Phe	260	265	270	275
Ser	Arg	Ser	Gln	Trp	Asp	Met	Gln	Ser	Thr	Phe	Asp	Asn	Thr	Ile	280	285	290	295
Lys	Glu	His	Phe	Pro	Glu	Gly	Ser	Arg	Pro	Val	Leu	Ser	His	Leu	300	305	310	315
Leu	Leu	Lys	Ala	Glu	Leu	Trp	Phe	Val	Asn	Ser	Asp	Phe	Ala	Phe	320	325	330	335
Asp	Phe	Ala	Arg	Pro	Leu	Leu	Pro	Asn	Thr	Val	Tyr	Ile	Gly	Gly	340	345	350	355
Leu	Met	Glu	Lys	Pro	Ile	Lys	Pro	Val	Pro	Gln	Asp	Leu	Asp	Asn	360	365	370	375
Phe	Ile	Ala	Asn	Phe	Gly	Asp	Ala	Gly	Phe	Val	Leu	Val	Ala	Phe	380	385	390	395
Gly	Ser	Met	Leu	Asn	Thr	His	Gln	Ser	Gln	Glu	Val	Leu	Lys	Lys	400	405	410	415
Met	His	Asn	Ala	Phe	Ala	His	Leu	Pro	Gln	Gly	Val	Ile	Trp	Thr	420	425	430	435
Cys	Gln	Ser	Ser	His	Trp	Pro	Arg	Asp	Val	His	Leu	Ala	Thr	Asn	440	445	450	455
Val	Lys	Ile	Val	Asp	Trp	Leu	Pro	Arg	Ser	Asp	Leu	Leu	Ala	His	460	465	470	475
Pro	Ser	Ile	Arg	Leu	Phe	Val	Thr	His	Gly	Gly	Gln	Asn	Ser	Val	480	485	490	495
Met	Glu	Ala	Ile	Arg	His	Gly	Val	Pro	Met	Val	Gly	Leu	Pro	Val	500	505	510	515
Asn	Gly	Asp	Gln	His	Gly	Asn	Met	Val	Arg	Val	Val	Ala	Lys	Asn	520	525	530	535
Tyr	Gly	Val	Ser	Ile	Arg	Leu	Asn	Gln	Val	Thr	Ala	Asp	Thr	Leu	540	545	550	555
Thr	Leu	Thr	Met	Lys	Gln	Val	Ile	Glu	Asp	Lys	Arg	Tyr	Lys	Ser	560	565	570	575
Ala	Val	Val	Ala	Ala	Ser	Val	Ile	Leu	His	Ser	Gln	Pro	Leu	Ser	580	585	590	595
Pro	Ala	Gln	Arg	Leu	Val	Gly	Trp	Ile	Asp	His	Ile	Leu	Gln	Thr	600	605	610	615

Gly	Gly	Ala	Thr	His	Leu	Lys	Pro	Tyr	Ala	Phe	Gln	Gln	Pro	Trp	
				470					475					480	
His	Glu	Gln	Tyr	Leu	Ile	Asp	Val	Phe	Val	Phe	Leu	Leu	Gly	Leu	
				485					490					495	
Thr	Leu	Gly	Thr	Met	Trp	Leu	Cys	Gly	Lys	Leu	Leu	Gly	Val	Val	
				500					505					510	
Ala	Arg	Trp	Leu	Arg	Gly	Ala	Arg	Lys	Val	Lys	Lys	Thr			
				515					520						

&lt;210&gt; 3

&lt;211&gt; 358

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7472775CD1

&lt;400&gt; 3

Met	Pro	Glu	Asp	Val	Arg	Glu	Lys	Lys	Glu	Asn	Leu	Leu	Leu	Asn	
1				5					10					15	
Ser	Glu	Arg	Ser	Thr	Arg	Leu	Leu	Thr	Lys	Thr	Ser	His	Ser	Gln	
				20					25					30	
Gly	Gly	Asp	Gln	Ala	Leu	Ser	Lys	Ser	Thr	Gly	Ser	Pro	Thr	Glu	
				35					40					45	
Lys	Leu	Ile	Glu	Lys	Arg	Gln	Gly	Ala	Lys	Thr	Val	Phe	Asn	Lys	
				50					55					60	
Phe	Ser	Asn	Met	Asn	Trp	Pro	Val	Asp	Ile	His	Pro	Leu	Asn	Lys	
				65					70					75	
Ser	Leu	Val	Lys	Asp	Asn	Lys	Trp	Lys	Lys	Thr	Glu	Glu	Thr	Gln	
				80					85					90	
Glu	Lys	Arg	Arg	Ser	Phe	Leu	Gln	Glu	Phe	Cys	Lys	Lys	Tyr	Gly	
				95					100					105	
Gly	Val	Ser	His	His	Gln	Ser	His	Leu	Phe	His	Thr	Val	Ser	Arg	
				110					115					120	
Ile	Tyr	Val	Glu	Asp	Lys	His	Lys	Ile	Leu	Tyr	Cys	Glu	Val	Pro	
				125					130					135	
Lys	Ala	Gly	Cys	Ser	Asn	Trp	Lys	Arg	Ile	Leu	Met	Val	Leu	Asn	
				140					145					150	
Gly	Leu	Ala	Ser	Ser	Ala	Tyr	Asn	Ile	Ser	His	Asn	Ala	Val	His	
				155					160					165	
Tyr	Gly	Lys	His	Leu	Lys	Lys	Leu	Asp	Ser	Phe	Asp	Leu	Lys	Gly	
				170					175					180	
Ile	Tyr	Thr	Arg	Leu	Asn	Thr	Tyr	Thr	Lys	Ala	Val	Phe	Val	Arg	
				185					190					195	
Asp	Pro	Met	Glu	Arg	Leu	Val	Ser	Ala	Phe	Arg	Asp	Lys	Phe	Glu	
				200					205					210	
His	Pro	Asn	Ser	Tyr	Tyr	His	Pro	Val	Phe	Gly	Lys	Ala	Ile	Ile	
				215					220					225	
Lys	Lys	Tyr	Arg	Pro	Asn	Ala	Cys	Glu	Glu	Ala	Leu	Ile	Asn	Gly	
				230					235					240	
Ser	Gly	Val	Lys	Phe	Lys	Glu	Phe	Ile	His	Tyr	Leu	Leu	Asp	Ser	
				245					250					255	
His	Arg	Pro	Val	Gly	Met	Asp	Ile	His	Trp	Glu	Lys	Val	Ser	Lys	
				260					265					270	
Leu	Cys	Tyr	Pro	Cys	Leu	Ile	Asn	Tyr	Asp	Phe	Val	Gly	Lys	Phe	
				275					280					285	
Glu	Thr	Leu	Glu	Glu	Asp	Ala	Asn	Tyr	Phe	Leu	Gln	Met	Ile	Gly	
				290					295					300	
Ala	Pro	Lys	Glu	Leu	Lys	Phe	Pro	Asn	Phe	Lys	Asp	Arg	His	Ser	
				305					310					315	
Ser	Asp	Glu	Arg	Thr	Asn	Ala	Gln	Val	Val	Arg	Gln	Tyr	Leu	Lys	
				320					325					330	